

Requested Patent: JP11024982A
Title: WEB PAGE LOOKAHEAD SYSTEM BASED ON HISTORY ;
Abstracted Patent: JP11024982 ;
Publication Date: 1999-01-29 ;
Inventor(s): YAMAMOTO KATSUYUKI ;
Applicant(s): NEC CORP ;
Application Number: JP19970189196 19970630 ;
Priority Number(s): ;
IPC Classification: G06F12/00 ;

Equivalents:

ABSTRACT:

PROBLEM TO BE SOLVED: To fast display the latest Web page data and to secure the excellent responsiveness of WWW browser software by holding a list of link destination in each Web page together with selection frequency based on past history and reading a page that has high possibility of being selected into a cache area in advance. **SOLUTION:** A table updating part 140 updates the selection frequency of a corresponding entry of a link table 170 when there is link selection. A gape lookahead part 150 performs lookahead of a link destination having high past selection frequency in the background based on the content of the table 170 and stores page data that are performed lookahead in page data cache 160. That is, it performs lookahead of a corresponding Web page in the background from a server about a link destination that has high past selection frequency in a link in which the displayed Web page is included and stores it in the cache 60. Therefore, the latest page data can be fast shown.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平11-24982

(43) 公開日 平成11年(1999) 1月29日

(51) Int.Cl.⁶

G 0 6 F 12/00

識別記号

5 4 7

F I

G 0 6 F 12/00

5 4 7 H

審査請求 有 請求項の数 5 F D (全 10 頁)

(21) 出願番号 特願平9-189196

(22) 出願日 平成9年(1997) 6月30日

(71) 出願人 000004237

日本電気株式会社

東京都港区芝五丁目7番1号

(72) 発明者 山本 勝之

東京都港区芝五丁目7番1号 日本電気株式会社内

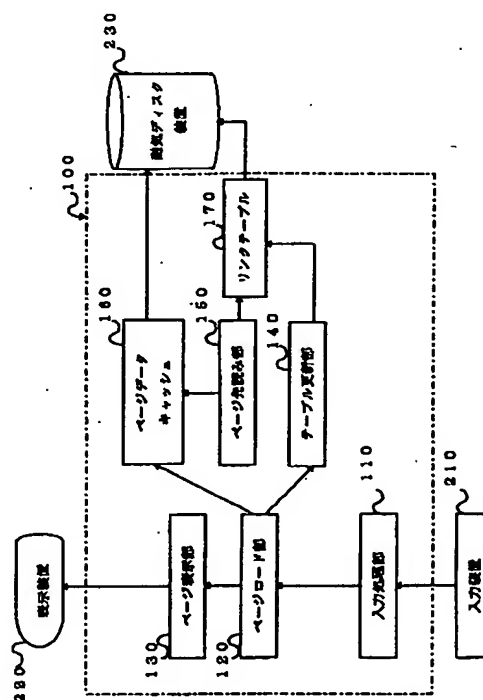
(74) 代理人 弁理士 加藤 朝道

(54) 【発明の名称】 履歴に基づくWebページ先読み方式

(57) 【要約】

【課題】最新のWebページデータの高速表示を可能とし、WWWブラウザソフトウェアの良好な応答性を確保するWeb先読み方式の提供。

【解決手段】表示装置および入力装置を有するコンピュータシステム上で動作し、サーバよりネットワーク経由で、Webページをダウンロードして表示するWWWブラウザにおいて、Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルと、リンクの選択が行われたときに前記リンクテーブルの内容を更新するテーブル更新部と、前記リンクテーブルの内容に基づき過去の選択頻度が高いURLのデータ先読みを行うページ先読み部と、を有し、選択される可能性が高いページを事前にキャッシュ領域に読み込む。



【特許請求の範囲】

【請求項1】表示装置および入力装置を有するコンピュータシステム上で動作し、サーバよりネットワーク経由で、Webページをダウンロードして表示するWWW (World Wide Web) ブラウザにおいて、Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルと、リンクの選択が行われたときに前記リンクテーブルの内容を更新するテーブル更新部と、前記リンクテーブルの内容に基づき過去の選択頻度が高いURL (Uniform Resource Locator) のデータ先読みを行うページ先読み部と、を有し、選択される可能性が高いページを事前にキャッシュ領域に読み込む、ことを特徴とするWebページ先読み方式。

【請求項2】WWW (World Wide Web) サーバとして動作するコンピュータシステム上で動作するWWWサーバソフトウェアシステムにおいて、Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルと、リンクの選択が行われたときに前記リンクテーブルの内容を更新するテーブル更新部と、前記リンクテーブルの内容に基づき過去の選択頻度が高いURL (Uniform Resource Locator) のデータ先読みを行うページ先読み部と、を有し、選択される可能性が高いページをクライアントコンピュータシステムのWWWブラウザが使用するキャッシュ領域に事前に転送する、ことを特徴とするWebページ先読み方式。

【請求項3】サーバよりネットワーク経由でWebページをダウンロードして表示するクライアント計算機のWWW (World Wide Web) ブラウザにおいて、Webページを表示するとリンクを解析し、該ページからのリンク一覧を生成し、リンクの選択が行われると、リンク一覧中の該当するリンクの選択頻度を上げる処理を行い、Webページにおけるリンクの選択頻度を蓄積し、既にリンク一覧を保持するWebページを再度表示した際、保持する前記リンク一覧から過去の選択頻度が高いリンクを検索し、検索されたリンク先のURL (Uniform Resource Locator) に対し、キャッシュに格納されるデータの調査および更新日時の比較を行い、前記キャッシュに該当データが存在しない、または前記キャッシュにあるデータの更新日時が古い場合、バックグラウンドで先読みを行って、前記キャッシュに格納し、

ユーザによるリンクの選択があった場合、次に表示するページデータが既にキャッシュに先読みされている可能性を高め、ユーザの平均待機時間を短縮する、ことを特徴とするWebページ先読み方法。

【請求項4】サーバよりネットワーク経由でWebページをダウンロードして表示するクライアント計算機のWWW (World Wide Web) ブラウザを構成する手段であって、

(a) Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルとの内容をリンクの選択が行われたときに更新するテーブル更新手段、

(b) 前記リンクテーブルの内容に基づき過去の選択頻度が高いURL (Uniform Resource Locator) のデータ先読みを行うページ先読み手段、
の上記各手段を前記計算機で機能させるためのプログラムを記録した記録媒体。

【請求項5】WWW (World Wide Web) サーバとして動作する計算機システム上で動作するWWWサーバソフトウェアシステムを構成する手段であって、

(a) Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルの内容をリンクの選択が行われたときに更新するテーブル更新手段、

(b) 前記リンクテーブルの内容に基づき過去の選択頻度が高いURL (Uniform Resource Locator) のデータ先読みを行うページ先読み手段、及び、

(c) 選択される可能性が高いページをクライアントコンピュータシステムのWWWブラウザが使用するキャッシュ領域に事前に転送する手段、
の上記各手段を前記計算機で機能させるためのプログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、履歴に基づくWebページ先読み方式に関し、特に過去のリンク選択履歴に基づいた選択頻度を保持するリンクテーブルを利用し、選択頻度が高いWebページのデータの先読みを可能にした履歴に基づくWebページ先読み方式に関する。

【0002】

【従来の技術】ディスプレイ等の表示装置およびキーボード、マウス等の入力装置を有するコンピュータシステム上で動作し、ネットワーク経由でWebページを表示するWWWブラウザソフトウェアにおいて、従来はユーザによりリンクの選択が行われた時点で、対象となるWebページを、ネットワーク経由で取得し、表示を行うものが一般的である。しかし、ユーザによるリンク選択の度、ネットワーク経由のトラフィックが発生し、そのためWWW (World Wide Web) ブラウザソフトウェア

がページ表示を完了するまでユーザの待機時間が発生する。クライアント端末におけるWWWブラウザソフトウェアにおいては、ユーザの待機時間が長くなることは運用面で大きな問題となる。

【0003】Webページ先読み方式は、上記問題の解消を図るため、Webページのようなハイパー文書の表示において、表示装置の応答性を良好にする目的で用いられている。

【0004】例えば特開平8-87526号公報には、ハイパー文書表示端末においてボタン/リンクが選択された履歴を保持し、ハイパー文書サーバの事前転送シーン判断部にて、前記履歴から得られる選択確率と、表示要求されたシーンからの距離を用いた関連度を算出し、より関連度が高いハイパー文書を、表示端末のメモリに事前に転送することによって、表示端末の応答性を良好にするようにしたハイパーメディア文書通信装置が提案されている。

【0005】

【発明が解決しようとする課題】しかし、この従来通信装置では、ハイパー文書サーバとハイパー文書表示装置の双方に構成要素が必要であり、インターネット上のWWWページのように、ハイパー文書の転送が特定のコンピュータシステム間に限られない場合においては、効果が期待できないという問題点があった。

【0006】また、従来の通信装置では、事前に転送されるハイパー文書が表示端末のメモリ上に置かれるため、過去に転送されたハイパー文書を再利用することができないという問題点も存在していた。

【0007】以上説明したように、上記従来技術は下記記載の問題点を有している。

【0008】第1の問題点は、インターネット上のWWWページを閲覧するためのWWWブラウザソフトウェアにおいては、接続するサーバを限定できないため、事前転送シーン判断部を保持しないサーバに接続した場合良好な応答性が得られない、ということである。

【0009】その理由は、従来の先読み方式では、ハイパー文書サーバとハイパー文書表示装置の双方に構成要素が必要であり、接続先のコンピュータシステムを限定できない環境についてなら考慮されていない、ためである。

【0010】第2の問題点は、サーバが事前転送したハイパー文書は転送直後に表示要求があった場合は良好な応答性が得られるが、その後の表示要求に対しては良好な応答性が期待できない、ということである。

【0011】その理由は、サーバが事前転送するハイパー文書は表示端末のメモリ上に配置されるため、転送済みのハイパー文書を以後再利用する方式について何ら考慮されていない、ためである。

【0012】したがって、本発明は上記問題点に鑑みてなされたものであって、その目的は、最新のWebページ

データの高速表示を可能とし、WWWブラウザソフトウェアの良好な応答性を確保したWeb先読み方式を提供することにある。

【0013】本発明の他の目的は、最新のWebページデータの高速表示を可能とし、クライアント側のWWWブラウザソフトウェアの良好な応答性を達成するWeb先読み方式を提供することにある。

【0014】

【課題を解決するための手段】前記目的を達成するため、本発明のWebページ先読み方式は、表示装置および入力装置を有するコンピュータシステム上で動作し、サーバよりネットワーク経由で、Webページをダウンロードして表示するWWW(World Wide Web)ブラウザにおいて、Webページ毎のリンク先一覧を、過去の履歴を基にした選択頻度とともに保持するリンクテーブルと、リンクの選択が行われたときに前記リンクテーブルの内容を更新するテーブル更新部と、前記リンクテーブルの内容に基づき過去の選択頻度が高いURL(Uniform Resource Locator)のデータ先読みを行うページ先読み部と、を有し、選択される可能性が高いページを事前にキャッシュ領域に読み込む、ことを特徴とする。

【0015】【発明の概要】本発明においては、Webページ毎のリンク先一覧と各リンク先に対する選択頻度をリンクテーブル(図1の170)に保持し、対象のWebページの表示が行われた場合参照する。より具体的には、Webページの表示が行われた場合リンクテーブル(図1の170)を参照し、過去の選択頻度が高いリンク先のWebページのデータをページ先読み部(図1の150)により先読みを行い、ページデータキャッシュ(図1の160)にダウンロードする。また、リンクの選択が行われ新たにWebページの表示が行われた場合、前記先読み処理を行うとともに、テーブル更新部(図1の140)がリンクテーブル(図1の170)内の選択されたリンクに対応するエントリの選択頻度情報を更新する。

【0016】

【発明の実施の形態】次に、本発明の実施の形態について図面を参照して詳細に説明する。

【0017】【実施の形態1】図1は、本発明の第1の実施の形態の構成を示すブロック図である。

【0018】図1を参照すると、本発明の第1の実施の形態は、マウス、キーボード等の入力装置210と、プログラム制御により動作するデータ処理装置100と、磁気ディスク等の記録装置230と、ディスプレイ等の表示装置220と、を含む。

【0019】またデータ処理装置100は、入力装置210からの入力を処理する入力処理部110と、Webページを構成するHTML(Hyper Text Markup Language)文書をロードするページロード部120と、ロードされたデータを画面に表示するページ表示部130

と、リンクの選択時に対応するエントリを更新するテーブル更新部140と、バックグラウンドでページデータを先読みするページ先読み部150と、先読みされたページデータのリストと共に格納するページデータキャッシュ160と、Webページ毎のリンク先リストを選択頻度と共に記憶するリンクテーブル170と、を含む。

【0020】図2は、リンクテーブル170の内容を示す図である。リンクテーブル170は、 n 個のWebページに対して $LT_1 \sim LT_n$ の n 個のテーブルを保持する。各テーブルは対応するWebページにおけるリンク先のURL (Uniform Resource Locator) を過去の選択頻度とともに保持する。

【0021】 LT_1 における対応WebページのURLは URL_0 に格納され、 m 個のリンク先のURLが $URL_1 \sim URL_m$ にそれぞれの過去の選択頻度 $f_1 \sim f_m$ と共に格納される。

【0022】図3は、ページデータキャッシュ160の内容を示す図である。ページデータキャッシュ160は、ページデータの実体を格納するキャッシュ領域と、キャッシュ領域に格納されているURLのリストを保持するキャッシュリストに分かれている。

【0023】キャッシュ領域は、過去ロードしたデータを保持する通常キャッシュ領域と、先読みを行ってロードされたデータを保持する先読みキャッシュ領域と、に分かれており、全キャッシュ領域のサイズを S 、通常キャッシュ領域のサイズを S_1 、先読みキャッシュ領域のサイズを S_2 としたとき、 $S = S_1 + S_2$ である。

【0024】キャッシュリストは、通常キャッシュ領域に格納されているページデータに対応するURLを k 個、先読みキャッシュ領域に格納されているページデータに対応するURLを1個、各ページデータの最終更新日時 $D_1 \sim D_k$ および $D_1 \sim D_1$ と共に保持する。

【0025】テーブル更新部140は、リンクの選択があった場合に、リンクテーブル170の該当するエントリの選択頻度を更新する。

【0026】ページ先読み部150は、リンクテーブル170の内容に基づいて過去の選択頻度が高いリンク先についてバックグラウンドで先読みを行い、ページデータキャッシュ160に先読みしたページデータを格納する。

【0027】本発明の第1の実施の形態では、ユーザがWWWブラウザソフトウェアによりWebページの表示を行った場合に、以下の処理を行う。

【0028】すなわち、表示しているWebページに含まれているリンクの中から、過去の選択頻度が高いリンク先について、バックグラウンドで該当するWebページをサーバから先読みし、キャッシュに格納する。

【0029】次に本発明の実施の形態の動作について説明する。

【0030】入力装置210による入力が発生したこと

が検出されると、データ処理部100が処理を開始する。

【0031】図4、図5、図6、及び図7は、データ処理部100における処理の流れを示すフローチャートである。図4は、本発明の実施の形態において、入力処理によりWebページの表示が行われた場合の処理を示す。

【0032】図1及び図4を参照すると、Webページの表示が行われた場合、はじめにリンクテーブル170内に該当するテーブルがあるかどうかを検査する(ステップ310)。ステップ310の判定の結果、テーブルが存在しなければ、リンクテーブル170内に新規テーブルを作成し(ステップ320)、テーブルを既に保持していた場合は、テーブル内の各リンクの選択頻度に基づいてページ先読み処理を行う(ステップ150)。

【0033】Webページ表示後、ページ内のリンクの選択が行われた場合には(ステップ330のYES分岐)、リンクテーブル170中の該当するリンク先の選択頻度を更新する(ステップ140)。

【0034】図5は、図4のステップ320のリンクテーブル作成処理の詳細な処理の流れを示すフローチャートである。

【0035】はじめに、対象ページ内に他のURLへのリンクが存在するかどうかを検査する(ステップ321)。リンクが一つも存在しなかった場合は処理を終了する。一方、リンクが存在した場合(ステップ321のYES分岐)、リンクテーブル170に新規テーブルを作成し、リンク先のURLのエントリを作成する(ステップ322)。

【0036】次にページ内の全リンクがテーブルに追加されたかどうかを検査し(ステップ323)、リンクが残っていれば次のリンクに処理を移す(ステップ324)。

【0037】図6は、図4のステップ140のリンクテーブル更新処理の詳細な処理の流れを示すフローチャートである。

【0038】はじめに選択されたリンクがリンクテーブル170のエントリに存在するかどうかを検査し(ステップ141)、存在しない場合、対象リンクのエントリをテーブルに追加する(ステップ142)。次に該当するリンクの選択頻度に1を加算する(ステップ143)。

【0039】図7は、図4のステップ150のページ先読み処理の詳細な処理の流れを示すフローチャートである。

【0040】はじめに対象リンクテーブル内で選択頻度が最も高いリンクを検索する(ステップ151)。

【0041】検索されたリンクについて、対象URLのページデータが既に通常キャッシュに存在するかどうかを検査し(ステップ152)、存在する場合はサーバ上

のデータと最終更新日時を比較する(ステップ153)。

【0042】通常キャッシュ内にページデータが存在しないか、またはキャッシュ内のデータがサーバ上のデータよりの古い場合は、対象URLのページデータが既に先読みキャッシュに存在するかどうかを検査し(ステップ154)、存在する場合はサーバ上のデータと最終更新日時を比較する(ステップ155)。

【0043】通常キャッシュ、先読みキャッシュいずれにもページデータが存在しないか、または、サーバ上のデータよりも古い場合には、ページデータキャッシュ160内の先読みキャッシュリストに、対象URLのエントリを追加し、先読みキャッシュ領域にページデータを読み込む(ステップ156)。

【0044】次にリンクテーブル170内の該当するテーブルの全てのリンクに対して処理が完了したかどうかを検査し(ステップ157)、処理完了した場合は終了し、未処理のリンクがある場合には、先読みキャッシュ領域に空き領域があるかどうかを検査し(ステップ158)、空き領域がある場合には、次に選択頻度が高いリンクを検索し(ステップ159)、ステップ152の通常キャッシュの検査に戻る。

【0045】

【実施例】上記した本発明の実施の形態について更に詳細に説明すべく、本発明の実施例について図8を参照して説明する。

【0046】図8は、あるリンク構造を持ったWebページの相関図である。はじめにページ1の表示が行われる(①)と、ページ1からリンクが張られているページ2、ページ3、ページ4に対する過去の選択頻度に応じ、先読みが行われる。

【0047】過去の選択頻度が、

[ページ2]>[ページ3]>[ページ4]

であったとすると、ページ1を表示しているバックグラウンドで、ページ2、ページ3、ページ4の順で先読みが行われ、各ページのデータがキャッシュ領域に格納される。

【0048】キャッシュ領域の空き領域が十分でなく、ページ2、ページ3のみ先読みしたとする。その場合も、次に選択されるリンクがリンク2またはリンク3である確率が高い。

【0049】次にリンク2を選択したとする(図8の②)と、通常キャッシュ内にページ1のデータが残るとともに、ページ2からのリンク先であるページ5のデータが先読みされる。

【0050】次にリンク5が選択されると(③)、ページ5からのリンク先であるページ1およびページ3のデータが選択頻度に応じた順序で先読みされるが、ページ1のデータは通常キャッシュに、ページ3のデータは先読みキャッシュに既に存在するため、無駄な先読み処理

が行われることはない。

【0051】次に、本発明の第2の実施の形態について図面を参照して詳細に説明する。

【0052】[実施の形態2]図9は、本発明の第2の実施の形態の構成を示すブロック図である。

【0053】図9を参照すると、本発明の第2の実施の形態は、前記第1の実施の形態のうち、ページ先読み部150、リンクテーブル170、テーブル更新部140を取り除き、これらの構成要素を、サーバ側に配設した点が相違している。

【0054】サーバ側では、クライアント側からWebページに対するアクセスがある度に、テーブル更新部430が対応するリンクテーブル420を更新し、選択頻度の高いリンクに対応するWebページを、ページ転送部410がクライアント側のページデータキャッシュ160に転送する。

【0055】以上の処理により、クライアント側に特別な構成要素を必要とせず、良好な応答性が期待できるという効果がある。

【0056】

【発明の効果】以上説明したように、本発明によれば下記記載の効果を奏する。

【0057】本発明の第1の効果は、WWWブラウザソフトウェアにより、あるWebページにおけるリンクを選択し、リンク先のWebページを表示する際、最新ページデータが高速に表示される、ことである。

【0058】その理由は、本発明においては、過去の選択頻度と最終更新日時を保持し、選択頻度が高いリンク先のページデータについて、操作者のページ閲覧中にバックグラウンドで最終更新日時の比較と更新を行い、最新データをキャッシュ領域に先読みすることにより、サーバからのダウンロードによる操作者の待機時間が発生しないためである。

【0059】本発明の第2の効果は、インターネット上のWWWページのようにサーバを限定しない環境であっても、第1の効果と同様の良好な応答性が期待できることである。

【0060】その理由は、本発明においては、クライアント側に本発明の先読み方式を組み込んでおくことにより、サーバ側に特別な構成要素を必要とせず効果的な先読みを行うことが可能なためである。

【0061】本発明の第3の効果は、クライアント側のコンピュータシステムの電源切断、再投入を経ても、電源切断前の先読み状態を保持し、再投入直後も良好な応答性が期待できることである。

【0062】その理由は、本発明においては、磁気ディスク装置などの外部記憶装置にキャッシュ領域を配置することにより、電源切断、再投入を経ても電源切断直前のキャッシュ領域状態が保持されるためである。

【図面の簡単な説明】

【図1】本発明の第1の実施の形態の構成を示すブロック図である。

【図2】本発明の第1の実施の形態におけるリンクテーブル170の内容の一例を示す図である。

【図3】本発明の第1の実施の形態におけるページデータキャッシュ160の内容の一例を示す図である。

【図4】本発明の第1の実施の形態のデータ処理部100の処理の流れを示すフローチャートである。

【図5】本発明の第1の実施の形態におけるリンクテーブル作成処理(図4の320)の処理の流れを示すフローチャートである。

【図6】本発明の第1の実施の形態におけるリンクテーブル更新処理(図4の140)の処理の流れを示すフローチャートである。

【図7】本発明の第1の実施の形態におけるページ先読み処理(図4の150)の処理の流れを示すフローチャートである。

ートである。

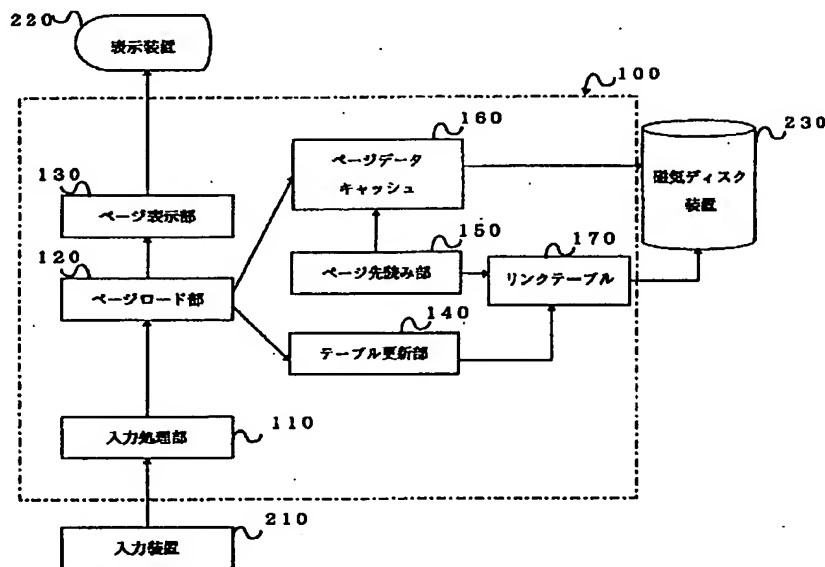
【図8】本発明による一実施例の動作を説明するためのWebページの相関図である。

【図9】本発明の第2の実施の形態の構成を示すブロック図である。

【符号の説明】

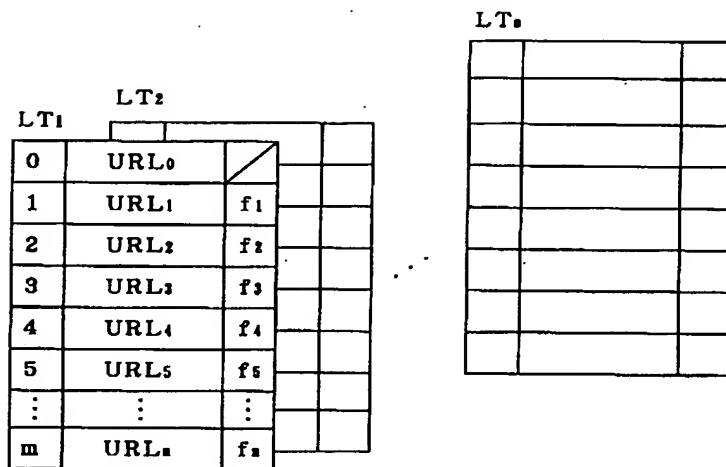
- 110 入力処理部
- 120 ページロード部
- 130 ページ表示部
- 140 テーブル更新部
- 150 ページ先読み部
- 160 ページデータキャッシュ
- 170 リンクテーブル
- 210 入力装置
- 220 表示装置
- 230 磁気ディスク装置

【図1】



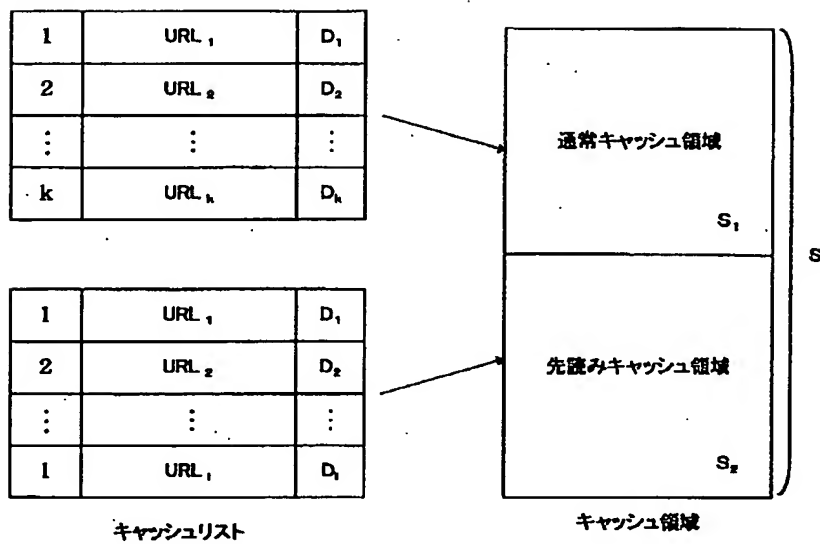
【図2】

170

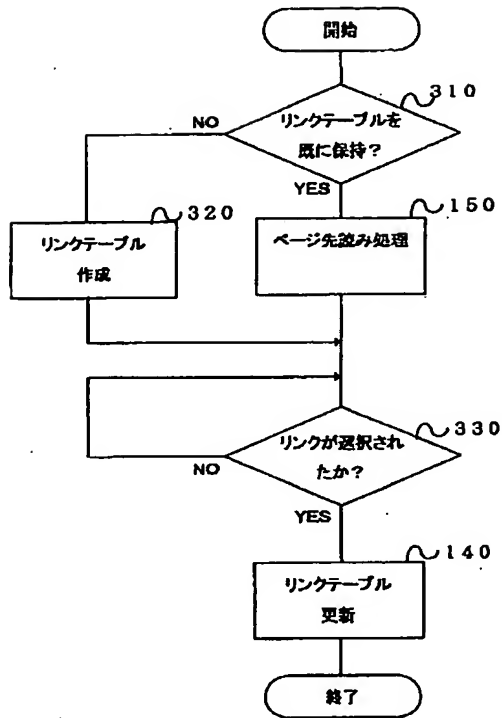


【図3】

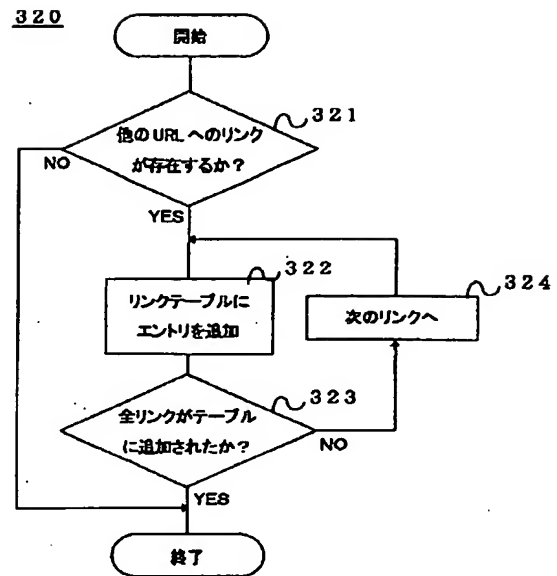
160



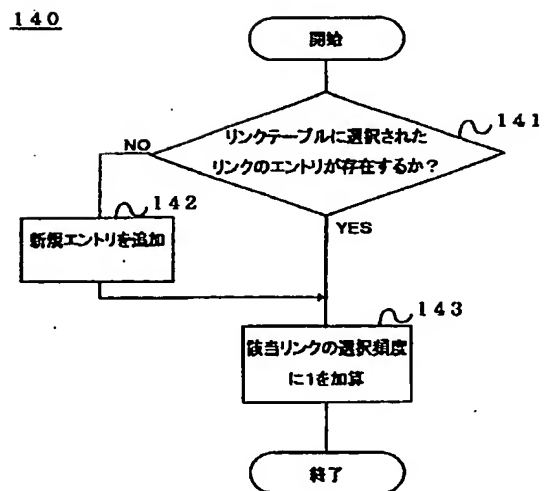
【図4】



【図5】

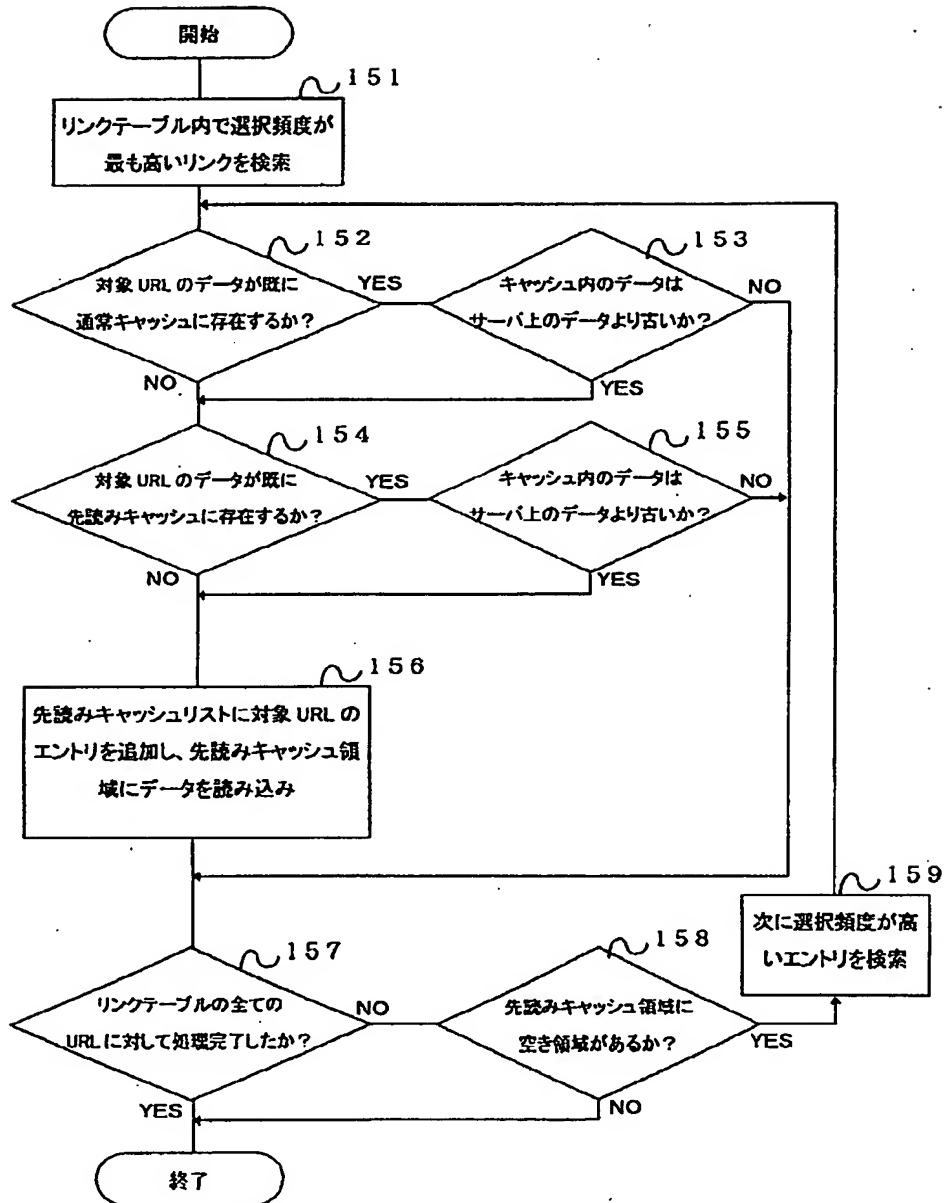


【図6】

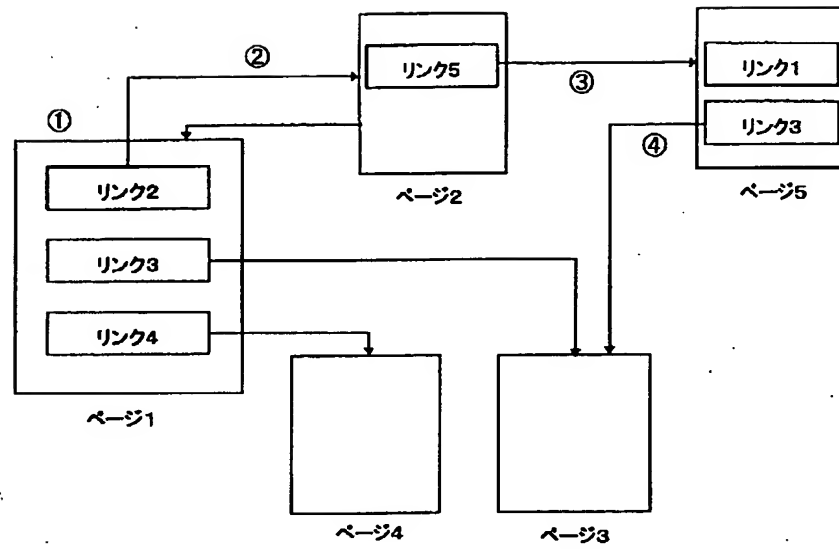


【図7】

150



【図8】



【図9】

